Interim Data Analysis Report: TellCo Analysis

1. Executive Summary

This interim report focuses on the progress of Task 1 under the TellCo project. The objective of Task 1 is to conduct a comprehensive user overview analysis by exploring customer behavior, device usage, and application activity. Key findings include insights into handset usage, session statistics, and application engagement metrics.

2. Introduction

The primary goal of Task 1 is to evaluate user behavior and preferences within the TellCo dataset. This includes identifying popular handsets, manufacturers, and analyzing user activity across various applications. The findings from this task will guide targeted marketing strategies and network optimization efforts.

3. Data Description

- Data Sources: The dataset was sourced from TellCo's telecom operations, focusing on Call Detail Records (CDR) and xDR data sessions.

- Dataset Overview: Key variables analyzed in this task include:

  - Customer_ID

  - Number_of_xDR_sessions

  - Session_Duration

  - Download_Data

  - Upload_Data

  - Handset_Type

  - Application_Usage

4. Methodology

- Tools Used: Python (pandas, matplotlib, seaborn).

- Analysis Steps:

- Aggregated user data to calculate total sessions, data usage, and session duration.
- Identified the top 10 handsets and top 3 manufacturers.
- Analyzed application engagement metrics such as download and upload data volumes.
- Segmented users based on total session duration into decile classes for further insights.

5. Preliminary Analysis

- Handset and Manufacturer Analysis:

  - Top 10 handsets: [List of top handsets].

  - Top 3 manufacturers: [List of manufacturers].

  - Insights: Manufacturer X dominates the market with Y% share of the top 10 handsets.

- User Behavior Metrics:

  - Average number of sessions: X.

  - Median session duration: Y minutes.

  - Total data volume (download + upload): Z TB.

- Application Usage:

  - Most engaged applications: Social Media, YouTube, and Gaming.

  - Insights: Social Media accounts for the highest data usage, followed by video streaming platforms.

6. Visualization

- Top Handsets: [Bar chart showing top 10 handsets].
- Application Engagement: [Pie chart showing data usage distribution among applications].
- User Segmentation: [Histogram of session duration decile classes].

7. Interim Findings

- The majority of users prefer handsets from Manufacturer X, indicating potential for partnerships or targeted marketing.
- Social Media and streaming applications dominate data usage, suggesting network optimization opportunities.
- User segmentation reveals high variability in session durations, with the top decile class accounting for Z% of total data usage.

8. Challenges Encountered

- Data Quality: Missing values in handset types required imputation using mode.
- Outliers: Extreme values in session duration and data usage were capped to improve analysis reliability.

9. Next Steps

- Expand analysis to include network performance metrics.
- Explore correlations between handset types and application preferences.
- Develop targeted recommendations for marketing and network resource allocation.

10. Conclusion

The analysis of Task 1 has provided valuable insights into user behavior and device preferences in TellCo's network. These findings will inform strategic decisions for enhancing customer satisfaction and optimizing service offerings. Further analysis in subsequent tasks will build on these insights to create a comprehensive understanding of user engagement and satisfaction.